

De los lexicones: NLP en la construcción del Lexicón de Drivers de Mercado en Español*

» **Alejandra M. J. Litterio**

Centro de Altos Estudios en Tecnología Informática (CAETI), Argentina
alejanda.litterio@UAI.edu.ar

Resumen

En nuestra aproximación sobre el análisis e interpretación de las noticias financieras sostenemos que, además de la necesidad de un lexicón de propósitos específicos para finanzas, es fundamental contar con lexicones de “drivers de mercado”, siendo un driver de mercado, aquel factor que ejerce un efecto material sobre una actividad de otra entidad, contextualmente dependiente y que afecta al mercado financiero en un momento determinado. Desde nuestro enfoque, proponemos: en primer lugar, conceptualizar las diferentes categorías de “drivers de mercado”, en segundo lugar, explicar de manera sucinta cómo mediante las técnicas de Procesamiento de Lenguaje Natural (NLP) se realiza la construcción del lexicón de drivers, y explicitar la relevancia del lexicón de drivers en la interpretación de noticias financieras y su correlación con los movimientos del mercado.

PALABRAS CLAVE: LEXICÓN- DRIVERS- MERCADOS FINANCIEROS- PROCESAMIENTO DE LENGUAJE NATURAL.

The lexicons: NLP in the design of Market Drivers Lexicon in Spanish

Abstract

In our approach when analyzing and interpreting financial news, we state that apart from the need of a tailored-made lexicon it is critical to design market drivers lexicons, defining a “driver” as the factor or force that has a material impact on a specific activity on another entity, which is contextually dependent and which affects the financial market at a specific time. Therefore,

* Este trabajo forma parte del Proyecto “Modelo de Sentiment Analysis para la clasificación de noticias en tiempo real” radicado en el Centro de Altos Estudios en Tecnología Informática – CAETI dirigido por el Dr. Alejandro Fernández y la Dra. Cristina Camós. Investigador y coordinador general: DS. Juan Pablo Braña.

our proposal contemplates three key aspects: first, the conceptualization of the market driver typology, second, a brief explanation of Natural Language Processing (NLP) techniques applied in the construction of market drivers lexicons to finally explain the relevance of the market drivers lexicons in the interpretation of financial news and its correlation to the market movements.

KEYWORDS: LEXICON-DRIVERS-FINANCIAL MARKETS-NATURAL LANGUAGE PROCESSING

1. Introducción

Habiendo realizado un recorrido exhaustivo de publicaciones en diferentes idiomas y si bien se ha constatado la existencia de trabajos pioneros en cuanto a la creación de lexicones de propósitos generales en diferentes idiomas (Moreno-Sandoval, 2017; Jiménez-Yáñez et. al., 2017; Shin et. al., 2017, Taboada et. al., 2011) así como lexicones que incluyen terminología específica en finanzas (Fersini et. al., 2014; Mao et. al., 2014; Wang et. al., 2013; Loughran and McDonald, 2011) es propio mencionar que hasta el momento, no se conocen trabajos en español que desde el Procesamiento de Lenguaje Natural den cuenta de la construcción, análisis composicional-semántico de lexicones de driver de mercado así como de la sistematización de las dinámicas enunciativas de las noticias económico-financiera intervinientes en el ámbito de la Bolsa y Mercado de Valores Argentinos (BYMA). Dicho de otro modo, no se cuenta con ningún antecedente de investigación específicamente dedicado al estudio de lexicones de drivers de mercado en español en cuanto a su funcionalidad desde el enfoque sistémico-funcional y el análisis de sentimiento como sub-campos del Procesamiento de Lenguaje Natural¹.

Ahora bien, generalmente, cuando los inversores se refieren a un mercado como *fundamentally driven* o *technically driven*² están utilizando una versión simplista, caracterizándolos de manera binaria. De hecho, los mercados constituyen un ecosistema donde millones de participantes que intervienen en las operaciones de *trading*³ lo hacen “movidos” por factores que trascienden estas operaciones y que orientan los movimientos mismos del mercado y donde los precios de los instrumentos financieros son un reflejo o una respuesta a eventos publicados en las noticias financieras.

La pregunta recae en comprender aquellas fuerzas o factores que mueven al mercado y que, como se observará en este trabajo, reúnen ciertas características dependiendo de las condiciones a nivel macro o micro y que ejercen un impacto en los mercados a nivel local o global. Así, un titular de una noticia como: “*Sube el riesgo país y se desploma la Bolsa porteña*” tiene un efecto altamente negativo en los mercados argentinos, donde el “riesgo

1 Como parte del proyecto de investigación en curso, nos encontramos en este momento desarrollando lexicones de drivers de mercado en inglés basados en las mismas concepciones teórico-metodológicas siguiendo la corriente de la Lingüística Sistémico-Funcional y las técnicas de Procesamiento de Lenguaje Natural.

2 Nos referimos en estos casos a los indicadores fundamentales y técnicos. Cuando un mercado es fundamentalmente impulsado, los precios de las acciones se relacionan estrechamente con los resultados de las operaciones de las empresas. Los factores fundamentales comunes incluyen ganancias por acción, ventas por acción, márgenes, flujo de caja, dividendos y la tasa de descuento. Definimos un mercado impulsado técnicamente como aquel en el que los precios están siendo fijados por cualquier cosa que no sea un indicador fundamental. Los indicadores técnicos más comunes son: Moving Average (MA), un indicador de tendencia, las bandas de Bollinger para medir la volatilidad y MACD, un indicador que mide la fuerza impulsora detrás del mercado.

3 Entiéndase como negociación. Aquí, la preferencia por el término en inglés resulta de su uso cotidiano en los mercados financieros argentinos.

país” no puede considerarse como un componente del lexicón de términos específicos sino que es, por definición, un driver, un factor, que merece ser aislado, y estudiado en términos de la creación de su propio lexicón con una terminología asociada. Lo mismo opera para el caso de: *“Las acciones locales mejoraron 15% desde el 9 de mayo, en parte impulsadas por el avance de negociaciones con el FMI, así como por la suba del dólar”*, donde el FMI y dólar se constituyen en drivers de mercado con un impacto a nivel local, de duración atemporal.

De esto se desprende que, siguiendo nuestra aproximación sobre el análisis e interpretación de las noticias financieras, sostengamos que contar con lexicones de “drivers de mercado” merece un tratamiento y análisis especial puesto que al ponerse en producción y en correlación con los movimientos del mercado, representan una diferencia importante en cuanto a la interpretación del evento y su repercusión con respecto al efecto que produce en la toma de decisiones tanto en una empresa como en el mercado financiero local o internacional. Recordemos que un hecho o evento considerado como “aislado” podría provocar un efecto contagio o epidémico⁴ en los mercados a nivel internacional.

2. Hacia una definición y categorización de drivers de mercado.

Para el caso particular de este trabajo, es preciso realizar una primera aproximación a la definición de “driver de mercado”.

En una acepción primigenia los drivers de mercados son “las fuerzas que obligan a los consumidores a comprar productos y pagar servicios”. Con respecto a la definición de drivers de mercado mencionada consideramos que es insuficiente para establecer un vínculo orientado a un contexto más amplio puesto que, en este caso, el foco analítico se encuentra en las variables o fuerzas que impulsan o buscan orientar el comportamiento de los jugadores en el mercado de consumo.

De modo que es necesario ampliar el espectro en busca de otras claves que respondan a aquellas formas que permitan a los jugadores del mercado reconocer los cambios, monitorear, explorar y utilizar la información del contexto para desarrollar nuevas estrategias.

Así, en el contexto bursátil, un driver de mercado es “un factor que ejerce un efecto material sobre una actividad de otra entidad, contextualmente dependiente y que afecta al mercado financiero en un momento determinado. Los denominados drivers se encuentran en los diferentes niveles de la economía⁵”. Dada esta definición, procedemos realizar una breve categorización de los drivers de mercado en la siguiente sección.

2.1. Tipología de Drivers.

La mayoría de los que trabajan en el sistema financiero, esto es, los administradores de activos, los economistas, los analistas y aquellos que tienen un interés particular en la economía global,

⁴ Si bien parece no haber un acuerdo ni en el campo académico ni en el campo aplicado para lo que se ha denominado contagio, el consenso parece moverse hacia una definición de contagio como la propagación de los tópicos o eventos en países de acuerdo a una dimensión temporal diseminados por las noticias financieras.

⁵ Mi traducción y reformulación a partir de la definición consultada en <https://www.investopedia.com/terms/d/driver.asp>

regularmente monitorean de cerca las publicaciones de datos económicos clave como el PIB, las cifras de desempleo y la inflación así como los indicadores técnicos y fundamentales que constituyen los principales factores que impulsan el desempeño económico de un país.

Cuantificar los datos estructurados así como los no estructurados, garantizando la coherencia y la precisión de las metodologías analíticas, es fundamental no sólo para medir el estado de una economía y el sistema financiero, sino también para llevar a cabo un análisis en tiempo real de aquellas fuerzas que producen un efecto en los mercados.

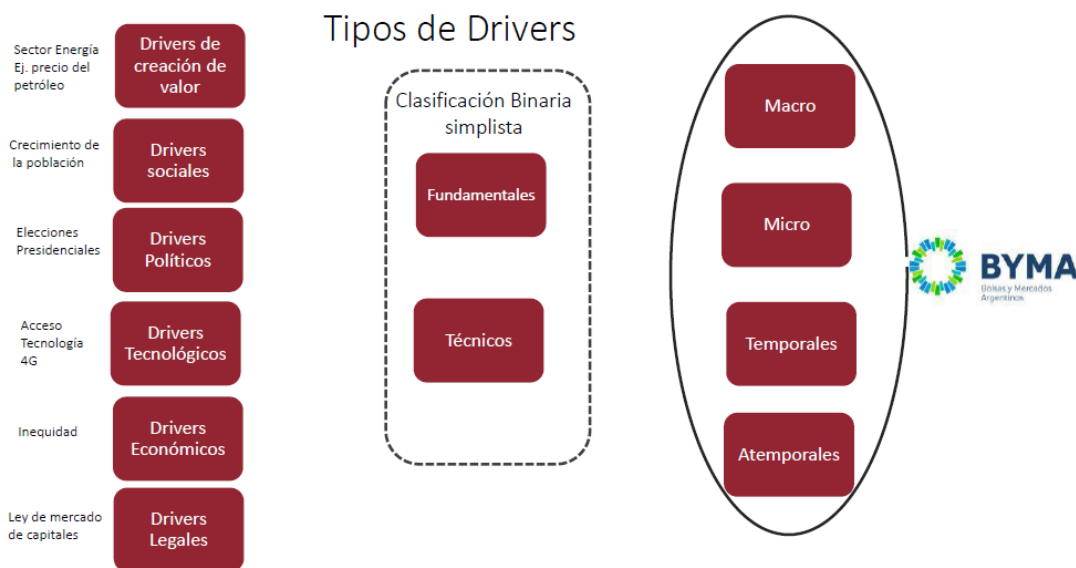


Figura 1. Tipología de Drivers

2.1.1. Macro, micro drivers y la temporalidad.

Es preciso realizar una primera distinción en cuanto a drivers de mercado entre macro y micro drivers. Mientras los primeros causan cambios a nivel global, los segundos provocan transformaciones a nivel local (véase Figura 2).

Los **macro drivers** afectan grandes sectores del mercado y generalmente incluyen eventos de gran importancia, como guerras, acuerdos comerciales u otros eventos geopolíticos.

Un **micro driver** es todo evento que pueda afectar materialmente a la economía y/o al sistema financiero de un país. Si hablamos a nivel empresa, algunos de los drivers más comunes incluyen el lanzamiento de nuevos productos o servicios, el nuevo financiamiento, los precios de los productos básicos o los recursos, las actividades de los competidores, la regulación y la diversificación de productos frente a los competidores.



Figura 2. Macro Drivers. Ejemplificación

Una segunda clasificación se deriva del efecto de temporalidad. Así, encontramos *drivers temporales y atemporales*, siendo los primeros transitorios, cuyo efecto en los mercados se suscribe a un período específico y no se prolongan en el tiempo, como por ejemplo la Ley de Mercado cambiario. Los segundos subsisten a través del tiempo, su efecto es duradero, como es el caso del Fondo Monetario Internacional (véase Figura 3).



Figura 3. Drivers temporales y atemporales

2.1.2. Drivers sociales, políticos, económicos y tecnológicos.

Un aspecto central de las dimensiones que vamos a caracterizar y ejemplificar de manera sucinta está relacionado con los factores que modelan los procesos de cambio y que afecta la estructura económica, social y política.

La siguiente clasificación es una síntesis de materiales bibliográficos consultados y que reúnen los rasgos centrales de cada caso:

- » Los **drivers sociales** se relacionan con las estructuras sociales e instituciones que determinan las preferencias, el comportamiento y las posibilidades de las personas, así como la autonomía, es decir, la capacidad de los individuos y los grupos para influir en el cambio. Así dentro de esta categoría podemos mencionar el crecimiento de la población, la sustentabilidad, el estilo de vida, etc.
- » Los **drivers políticos** tienen que ver con eventos en el espacio internacional como por ejemplo crisis mundiales (la crisis en España⁶ o en Italia⁷), conflictos geopolíticos o segregacionistas (las elecciones presidenciales, el Brexit⁸).
- » Los **drivers económicos** son aquellos factores o “inductores” que afectan el desarrollo económico a nivel local pero que pueden ejercer un impacto global. Ejemplo de ellos son la inflación, el cambio en las políticas monetarias, las políticas impositivas, la desigualdad económica, el desempleo, entre otros.
- » Los **drivers tecnológicos** son aquellas tendencias que marcan el avance y el uso de las diferentes tecnologías en situaciones de la vida cotidiana. Entre ellos se pueden enumerar: los sistemas localizadores de posicionamiento global (GPS), las tecnologías móviles (celulares, apps, etc.), agentes para la interacción personal (correo electrónico, mensajería instantánea, y similares), para las transacciones financieras (pago de facturas, de acciones, etc.), para la recopilación de información (por ejemplo, búsquedas en Internet, redes sociales).

Todos estos drivers denominados, por diferentes autores, como indicadores, factores, impulsores, constituyen un eje central en nuestra investigación puesto que ejercen una influencia en los mercados financieros locales e internacionales y, por lo tanto, merecen un tratamiento separado y la construcción de un lexicón (para cada driver) que refleje de manera fiel los cambios que se producen efectivamente en cuanto a la correlación entre la polaridad positiva-negativa de un evento micro o macro económico en las noticias financieras y las decisiones del mercado.

3. Lexicón de Drivers de Mercado

En este apartado, y dada la extensión del presente trabajo, describiremos brevemente la estructura composicional de los lexicones de drivers de mercado (de ahora en adelante LD) y brindaremos algunos casos a título ejemplificativo.

6 Cataluña y la triple crisis española. Consultado en: <http://nuso.org/articulo/cataluna-y-la-triple-crisis-espanola/>

7 La crisis en Italia hundió las bolsas de todo el mundo: el Dow Jones perdió 1,58% y el euro cayó ante el dólar a su menor valor en 10 meses. Consultado en:

<https://www.infobae.com/america/mundo/2018/05/29/la-crisis-en-italia-dispara-la-prima-de-riesgo-y-hunde-las-bolsas-de-europa/>

8 Qué es el Brexit y cómo puede afectar a Reino Unido y a la Unión Europea. Consultado en:

<https://www.bbc.com/mundo/noticias-internacional-36484790>

Para dar continuidad a nuestra línea de trabajo en cuanto a la construcción de lexicones de dominio específico en finanzas para los mercados argentinos (FSAL), decidimos aplicar la misma aproximación teórico-metodológica en el diseño de lexicones de drivers de mercado.

En este sentido diremos que para cada driver de mercado es preciso modelar un lexicón cuyas características se basan en la selección léxica de aquellos componentes que poseen un contenido semántico, esto es sustantivos, adjetivos, verbos y adverbios y donde se excluyen aquellos términos que por definición son “vacíos” o que no tienen una carga semántica, es decir, las preposiciones, los artículos, los demostrativos entre otros⁹. Además, cabe resaltar que, como un elemento más de la conformación del lexicón, hemos resuelto asignar a los términos una escala de apreciación /valoración en un rango de [-5, 5], siendo efectivamente [-5] el más negativo y [5] el más positivo.

Veamos algunos ejemplos tomados de titulares de noticias financieras en tiempo real.

(1) *Sube el Riesgo País y se desploma la Bolsa porteña.*

En el caso (1) desde el análisis semántico se observa que es un enunciado negativo representado por el hecho de que la suba del riesgo país produce un efecto rebote en la Bolsa porteña que cae abruptamente. Si el estudio se realizara siguiendo la metodología donde el término *riesgo* formara parte del FSAL entonces claramente se le asignaría un valor negativo [-5] y “sube” tendría un valor positivo [+5]. Ahora bien, si tomamos *riesgo país* como un indicador o driver de mercado y no como parte del FSAL, el componente “sube” tendría un valor negativo [-5], puesto que este término estaría asociado directamente al driver. Aquí es donde radica la importancia de la recomposición y construcción de un nuevo lexicón.

Observemos la composición vectorial desde NLP, donde el gradiente comprende una valoración muy positiva (VP), positiva (P), negativa (N) y muy negativa (VN) y donde Tw circunscribe al enunciado (evento) contenido en el tweet en un caso para el lexicón de términos específicos en finanzas (FSAL) y en el otro para el lexicón de driver de mercado (LD):

	VP	P	N	VN	
FSAL	0	1	1	1	Tw→ Neg
LD	0	0	1	1	Tw→ Neg

Nótese que, si bien no se presenta una diferencia importante a nivel categorización del titular de la noticia representado en el tweet ya que como se puede ver en el vector en ambos casos la tendencia es negativa, lo que resulta importante mencionar, en una primera instancia, es el cambio en cuanto a la valoración asignada al término, por ejemplo “sube”. Cuando este término es analizado como parte del lexicón FSAL, se le asigna una valoración positiva, que es modificada por una valencia negativa cuando el mismo es considerado un componente del lexicón de drivers (LD). Este primer hallazgo, es un reflejo sintomático-alerta que nos permitió realizar

⁹ Este tema y su descripción fue tratado en detalle en Juan Pablo Braña, Alejandra M. J. Litterio, Alejandro Fernández (2018): “FSAL: Lexicón financiero de sentimiento en español rioplatense diseñado para Bolsas y Mercados Argentinos (BYMA)”. Revista Abierta de Informática Aplicada, vol. 2 (1), pp. 5-22, ISSN 2591-5320.

nuevas pruebas con distintos enunciados de nuestro corpus y validar nuestra hipótesis: Un mismo término asociado a distintos lexicones cambia el sentido del enunciado y, por ende, su valoración o tendencia final.

En otras palabras, un cambio en la valorativo ínfimo puede ejercer una modificación total en su estructura interna sin producir una alteración drástica del resultado como lo refleja el término “sube” en el caso (1) donde el cambio es interno: de una valoración positiva a una negativa y sin embargo no se altera el resultado: el tweet es negativo tanto si forma parte del lexicón FSAL como si fuera un componente del lexicón de drivers (LD).

En otros casos, hay otros componentes en los que no sólo se altera la composición interna sino que el cambio en el resultado es radical: un tweet que se clasificó en un primer momento como positivo cuando el término es asociado al FSAL se transforma en negativo cuando el mismo término es asociado al lexicón de drivers, como se muestra en el caso (2) a continuación. Aquí se observa en mayor detalle cómo opera el lexicón de drivers de mercado:

(2) *El Riesgo País baja y la Bolsa trae esperanzas a los inversores.*

Tomando el ejemplo (2) podemos realizar el mismo estudio contrastivo. En una primera aproximación, si incluimos “riesgo” como componente semántico del lexicón de terminología específica (FSAL), entonces se debe asignar una valoración negativa [-5] y en el caso de los demás elementos constitutivos del lexicón FSAL, la distribución de la carga valorativa sería distribuida de la siguiente manera: baja [-3] y esperanzas [+4] (véase Figura 4). Si por el contrario tomamos “riesgo país” como un driver de mercado, el cual tiene asociado su propio lexicón, entonces la composición en cuanto a la valoración asignada a cada elemento varía: baja [+3] y esperanzas [+4].

Es relevante mencionar que, cuando se realiza el análisis aplicando técnicas de NLP, entonces en el vector se produce un cambio cualitativo importante alterándose por completo la lectura de semántica del titular de la noticia expresado en el tweet. Cuando se contempla “riesgo” como un elemento constitutivo del FSAL el tweet se clasifica como negativo, sin embargo si consideramos, tal y como se mencionaba en el ejemplo (2), “riesgo país” como un driver de mercado, entonces el tweet se categoriza como positivo¹⁰. Es precisamente aquí donde radica la esencia de poder realizar las revisiones y modificaciones correspondientes al trabajo anterior sobre lexicones de terminología específica para evitar una clasificación errónea pero más aún una interpretación equivocada de la noticia o de un tweet o comentario y por ende, incurrir en un error al generar el modelo de entrenamiento del algoritmo con inteligencia artificial. Recordemos que, en el caso que nos ocupa, el algoritmo “aprende” emulando el comportamiento humano.

¹⁰ Por otra parte, una pregunta que surgió a partir del estudio de drivers de mercado tiene un eje doble: ¿Qué sucede cuando un driver es afectado por otro driver o dos drivers o más se conjugan en el mismo evento afectando los movimientos del mercado?

Caso 1: “El riesgo país cae con fuerza en medio de negociaciones con el FMI”.

Caso 2: Suben las acciones argentinas por el acercamiento del FMI y La Ley de Financiamiento Productivo”

Con respecto a (1): siendo D1: riesgo país y D2: FMI, entonces ¿podría decirse que D2 afecta a D1?, y de ser así ¿cuál sería su tratamiento basados en técnicas de NLP? En cuanto a (2), siendo D1: FMI un macro driver atemporal y D2: Ley de Financiamiento un micro driver temporal, ¿cómo operan estas fuerzas en los movimientos del mercado?

Estos nuevos interrogantes son parte de un estudio en progreso como parte del proyecto de Trading Algorítmico y Sentiment Análisis para los mercados argentinos.

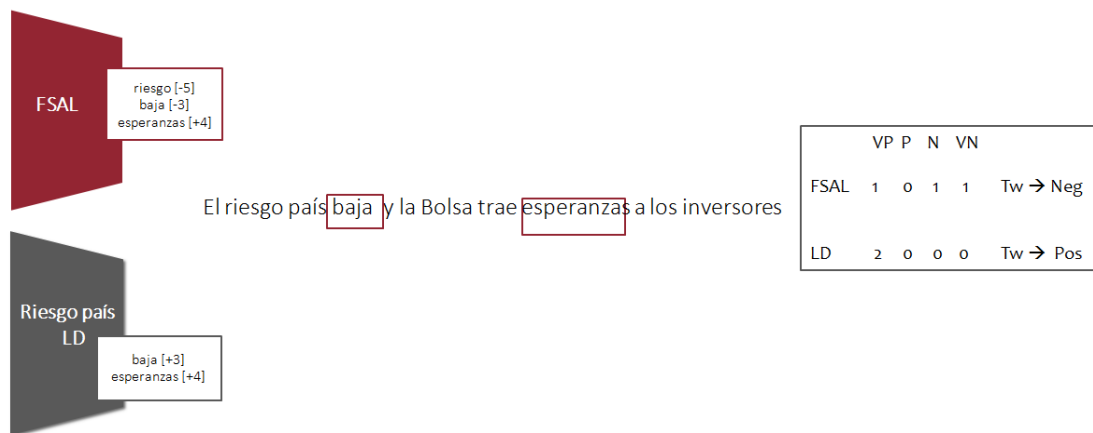
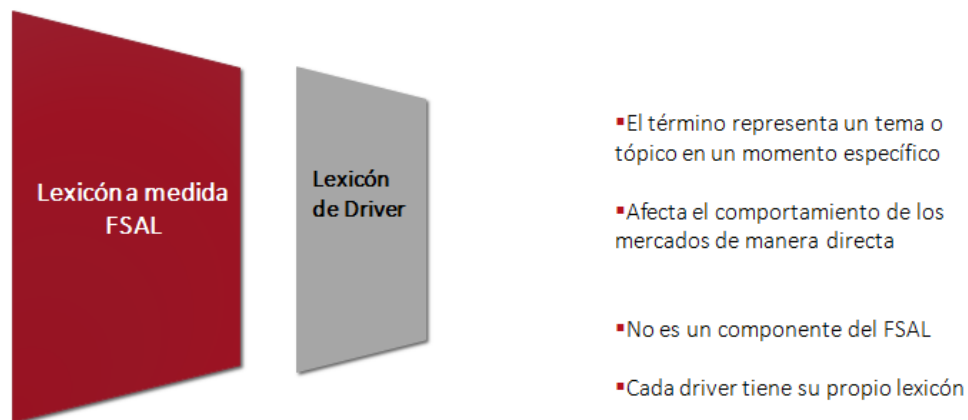


Figura 4. Análisis contrastivo de lexicones FSAL y LD

Es preciso destacar que no existe un único lexicón de driver de mercado, sino que para cada driver –inflación, dólar, FMI, petróleo, crisis, riesgo país, etc.– se debe construir un lexicón con la terminología asociada.

Resumiendo:



Finalmente, veamos cómo se articulan los lexicones en nuestro NLP Framework.

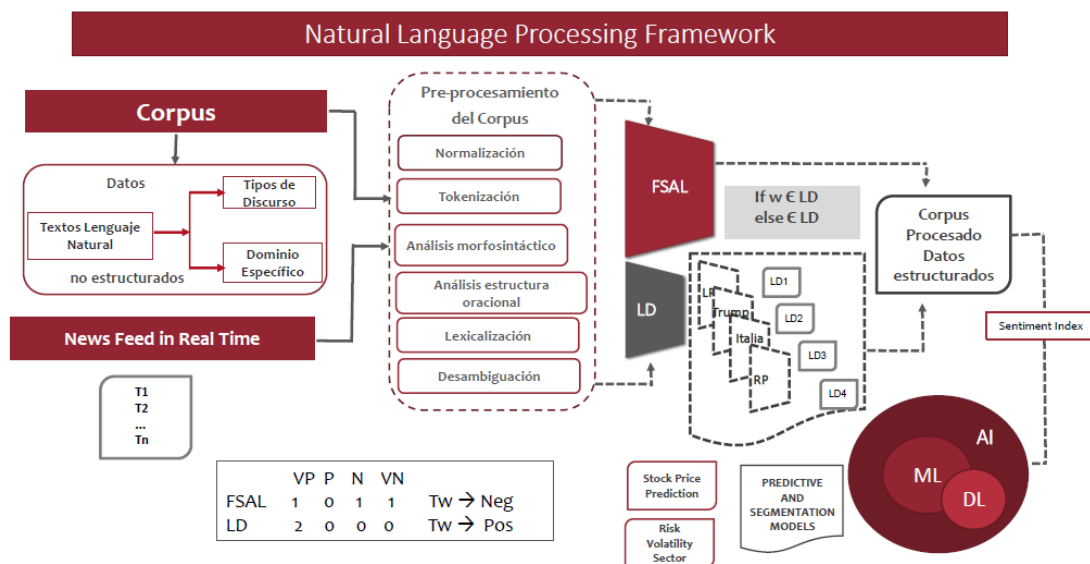


Figura 6. NLP Framework y los Lexicones

Como se observa en la Figura 6, nuestro corpus de análisis está compuesto de datos no estructurados, esto es, textos en lenguaje natural que corresponden a un tipo de discurso: las noticias en tiempo real del ámbito de las finanzas y economía. Una vez que estos datos han sido procesados, en otras palabras, han sido sometidos a diversas técnicas como la normalización, la descomposición en unidades mínimas de análisis (tokenización), la asignación de etiquetas a nivel morfosintáctico y el estudio de la estructura oracional (descomposición en unidades a nivel sintagmático y paradigmático), además de identificar el sentido en que la palabra se utiliza en una oración (desambiguación), se transforman en datos estructurados plausibles de ser analizados. De modo que si el término alojado en la estructura se encuentra como componente del lexicón de driver de mercado o en el lexicón FSAL, según sea el caso, entonces aplicando herramientas de *Machine Learning* podemos diseñar un índice de sentimiento lo que a su vez nos permite crear modelos predictivos y de segmentación. En el primer caso, el modelo podrá predecir la compra o venta de un instrumento financiero y establecer una ulterior correlación con los movimientos del mercado. En el segundo caso, se trata de crear un modelo de segmentación contemplando variables como el riesgo, la volatilidad y los sectores para detectar instrumentos que posean un perfil o comportamiento similar.

4. Comentarios finales

Como punto central de este trabajo se ha mostrado la necesidad de la creación de un nuevo lexicón y la aplicación de una nueva metodología de hibridación: FSAL + LD.

Hemos presentado la operatoria de los mecanismos que se activan cuando se utiliza un término asociado al FSAL y al LD, respectivamente. En otras palabras, los enunciados eventivos exhiben un status epistémico diferente cuando el término afectado pertenece a uno u otro lexicón.

Finalmente, acordamos que la creación y utilización de un lexicón de driver brinda una mejor posibilidad para redefinir con mayor grado de precisión la tendencia de un evento (en el caso

que nos ocupa una noticia financiera representada en un tweet) para la toma de decisiones por parte del algoritmo con Inteligencia Artificial, ya sea en la compra o venta de un instrumento financiero o para el rebalanceo de portfolios.

Referencias

- » Biwer, T., Jackobsen, B., and Kurkiewicz, A. (2012). Market Drivers: Fundamentals vs. Technicals. Is the market driven by fundamental factors, technical factors, or a mix of both? *FCA Magazine*, June 2012.
- » Fersini E., Messina E., Pozzi FA. (2014). Automatic construction of financial semantic orientation lexicon from large scale Chinese news corpus. *Decis Support Syst*, (68), 26–38.
- » Jimenez-Yañez, R., Gonzalez-Sanjurjo, H. Rayson, P., and Piao, S. (2017). Building a Spanish lexicon for corpus analysis. Conference: 35th Edition of the International Conference of The Spanish Association of Applied Linguistics (AESLA). Jaén, Spain Volume: *Languages at the Crossroads: Training, Accreditation and Context of Use*, 227-339. ISBN: 978-84-9159-108-5
- » Loughran T, McDonald B. (2011). When is a liability not a liability? Textual analysis, dictionaries. *J Finance*, (66), 35–65.
- » Mao H, Gao P, Wang Y, Bollen J. (2014). Automatic construction of financial semantic orientation lexicon from large scale Chinese news corpus. *7th Financial risks international forum*.
- » Mitnik, S., Robinzonov, N. and Spindler M. (2015). Stock market volatility: Identifying major drivers and the nature of their impact. *Journal of Banking and Finance*, (58), 1–14.
- » Moreno-Sandoval, L., Beltrán-Herrera, P., Vargas-Cruz, J., Sánchez-Barriga, C., Pomares-Quimbaya, A., Alvarado-Valencia, J. and C. García-Díaz, J. (2017). CSL: A Combined Spanish Lexicon Resource for Polarity Classification and Sentiment Analysis. In *Proceedings of the 19th International Conference on Enterprise Information Systems (ICEIS 2017)*, (1), 288-295.
- » Shin, B., Lee, T. and Choi, J. (2017). Lexicon Integrated CNN Models with Attention for Sentiment Analysis. *Proceedings of the 8th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*, 149–158. Copenhagen, Denmark, September 7–11, 2017
- » Song, G. (2018). The drivers of the Great Bull Stock Market of 2015 in China: Evidence and Policy Implications.
- » Taboada, M., Brooke, J., Tofiloski, M., Voll, K. and M. Stede (2011) "Lexicon-Based Methods for Sentiment Analysis". *Association for Computational Linguistics*, vol. 37, (2), 267-307.

